

0590
2/02

6 OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/971,020DATE: 01/02/2002
TIME: 11:08:40Input Set : A:\026350-068.ST25.txt
Output Set: N:\CRF3\01022002\I971020.raw**ENTERED**

3 <110> APPLICANT: Sano, Hiroshi
 4 Kusano, Tomonobu
 5 Koizumi, Nozomu
 7 <120> TITLE OF INVENTION: Theobromine Synthase Polypeptide of Coffee Plant and the

Gene

8 Encoding Said Polypeptide

10 <130> FILE REFERENCE: 026350-068

12 <140> CURRENT APPLICATION NUMBER: US 09/971,020

C--> 13 <141> CURRENT FILING DATE: 2001-12-18

15 <150> PRIOR APPLICATION NUMBER: JP 2000-307,149

16 <151> PRIOR FILING DATE: 2000-10-06

18 <160> NUMBER OF SEQ ID NOS: 8

20 <170> SOFTWARE: FastSEQ for Windows Version 4.0

22 <210> SEQ ID NO: 1

23 <211> LENGTH: 378

24 <212> TYPE: PRT

25 <213> ORGANISM: Coffea arabica

27 <400> SEQUENCE: 1

28 Met Glu Leu Gln Glu Val Leu His Met Asn Glu Gly Glu Gly Asp Thr
 29 1 5 10 15
 30 Ser Tyr Ala Lys Asn Ala Ser Tyr Asn Leu Ala Leu Ala Lys Val Lys
 31 20 25 30
 32 Pro Phe Leu Glu Gln Cys Ile Arg Glu Leu Leu Arg Ala Asn Leu Pro
 33 35 40 45
 34 Asn Ile Asn Lys Cys Ile Lys Val Ala Asp Leu Gly Cys Ala Ser Gly
 35 50 55 60
 36 Pro Asn Thr Leu Leu Thr Val Arg Asp Ile Val Gln Ser Ile Asp Lys
 37 65 70 75 80
 38 Val Gly Gln Glu Glu Lys Asn Glu Leu Glu Arg Pro Thr Ile Gln Ile
 39 85 90 95
 40 Phe Leu Asn Asp Leu Phe Gln Asn Asp Phe Asn Ser Val Phe Lys Leu
 41 100 105 110
 42 Leu Pro Ser Phe Tyr Arg Lys Leu Glu Lys Glu Asn Gly Arg Lys Ile
 43 115 120 125
 44 Gly Ser Cys Leu Ile Ser Ala Met Pro Gly Ser Phe Tyr Gly Arg Leu
 45 130 135 140
 46 Phe Pro Glu Glu Ser Met His Phe Leu His Ser Cys Tyr Ser Val His
 47 145 150 155 160
 48 Trp Leu Ser Gln Val Pro Ser Gly Leu Val Ile Glu Leu Gly Ile Gly
 49 165 170 175
 50 Ala Asn Lys Gly Ser Ile Tyr Ser Ser Lys Gly Cys Arg Pro Pro Val
 51 180 185 190
 52 Gln Lys Ala Tyr Leu Asp Gln Phe Thr Lys Asp Phe Thr Thr Phe Leu
 53 195 200 205
 54 Arg Ile His Ser Lys Glu Leu Phe Ser Arg Gly Arg Met Leu Leu Thr
 55 210 215 220
 56 Cys Ile Cys Lys Val Asp Glu Phe Asp Glu Pro Asn Pro Leu Asp Leu
 57 225 230 235 240

RAW SEQUENCE LISTING

DATE: 01/02/2002

PATENT APPLICATION: US/09/971,020

TIME: 11:08:40

Input Set : A:\026350-068.ST25.txt

Output Set: N:\CRF3\01022002\I971020.raw

```

58 Leu Asp Met Ala Ile Asn Asp Leu Ile Val Glu Gly Leu Leu Glu Glu
59                245                250                255
60 Glu Lys Leu Asp Ser Phe Asn Ile Pro Phe Phe Thr Pro Ser Ala Glu
61                260                265                270
62 Glu Val Lys Cys Ile Val Glu Glu Glu Gly Ser Cys Glu Ile Leu Tyr
63                275                280                285
64 Leu Glu Thr Phe Lys Ala His Tyr Asp Ala Ala Phe Ser Ile Asp Asp
65                290                295                300
66 Asp Tyr Pro Val Arg Ser His Glu Gln Ile Lys Ala Glu Tyr Val Ala
67 305                310                315                320
68 Ser Leu Ile Arg Ser Val Tyr Glu Pro Ile Leu Ala Ser His Phe Gly
69                325                330                335
70 Glu Ala Ile Met Pro Asp Leu Phe His Arg Leu Ala Lys His Ala Ala
71                340                345                350
72 Lys Val Leu His Met Gly Lys Gly Cys Tyr Asn Asn Leu Ile Ile Ser
73                355                360                365
74 Leu Ala Lys Lys Pro Glu Lys Ser Asp Val
75                370                375
78 <210> SEQ ID NO: 2
79 <211> LENGTH: 1298
80 <212> TYPE: DNA
81 <213> ORGANISM: Coffea arabica
83 <400> SEQUENCE: 2
84 agcagtcgca attcgattgt cctgcatatg aatggagctc caagaagtcc tgcataatgaa 60
85 tgaaggtgaa ggcgatacaa gctacgccaa gaatgcatcc tacaatctgg ctcttgccaa 120
86 ggtgaaacct ttccttgaac aatgcatacg agaattgttg cgggccaact tgcccaacat 180
87 caacaagtgc attaaagttg cggatttggg atgcgcttct ggaccaaaca cacttttaac 240
88 agtgcgggac attgtgcaaa gtattgacaa agttggccag gaagagaaga atgaattaga 300
89 acgtcccacc attcagattt ttctgaatga tcttttccaa aatgatttca attcggtttt 360
90 caagtgtctg ccaagcttct accgcaaaact cgagaaagaa aatggacgca agataggatc 420
91 gtgcctaata agcgcaatgc ctggctcttt ctacggcaga ctcttccccg aggagtccat 480
92 gcattttttg cactcttgtt acagtgttca ttggttatct caggttccca gcggtttggt 540
93 gattgaattg gggattggtg caaacaagg gagtatttac tcttccaaag gatgtcgtcc 600
94 gcccgctccag aaggcatatt tggatcaatt tacgaaagat tttaccacat ttctaaggat 660
95 tcattcgaaa gagttgtttt cacgtggccg aatgctcctt acctgcattt gtaaagtaga 720
96 tgaattcgac gaaccgaatc ccctagactt acttgacatg gcaataaacg acttgattgt 780
97 tgagggactt ctggaggaag aaaaattgga tagtttcaat attccattct ttacaccttc 840
98 agcagaagaa gtaaagtgca tagttgagga ggaaggttct tgcgaaattt tataatctgga 900
99 gacttttaag gccattatg atgctgcctt ctctattgat gatgattacc cagtaagatc 960
100 ccatgaacaa attaaagcag agtatgtggc atcattaatt agatcagttt acgaacccat 1020
101 cctcgcaagt cattttggag aagctattat gcctgactta ttccacaggc ttgcgaagca 1080
102 tgacgcaag gttctccaca tgggcaagg ctgctataat aatcttatca tttctctcgc 1140
103 caaaaagcca gagaagtcag acgtgtaaaa gtttgttttt agttggtttt tgtgccgttg 1200
104 ggggtctttc ggtattgtc gttttgtatt cgtaataaaa gtgatgtgca agaataagat 1260
105 atttagtaca atattttcat aaaaaaaaaa aaaaaaaaaa 1298
107 <210> SEQ ID NO: 3
108 <211> LENGTH: 385
109 <212> TYPE: PRT
110 <213> ORGANISM: Coffea arabica

```

RAW SEQUENCE LISTING

DATE: 01/02/2002

PATENT APPLICATION: US/09/971,020

TIME: 11:08:40

Input Set : A:\026350-068.ST25.txt

Output Set: N:\CRF3\01022002\I971020.raw

```

112 <400> SEQUENCE: 3
113 Met Glu Leu Gln Glu Val Leu His Met Asn Gly Gly Glu Gly Glu Ala
114 1 5 10 15
115 Ser Tyr Ala Lys Asn Ser Ser Phe Asn Gln Leu Val Leu Ala Lys Val
116 20 25 30
117 Lys Pro Val Leu Glu Gln Cys Val Arg Glu Leu Leu Arg Ala Asn Leu
118 35 40 45
119 Pro Asn Ile Asn Lys Cys Ile Lys Val Ala Asp Leu Gly Cys Ala Ser
120 50 55 60
121 Gly Pro Asn Thr Leu Leu Thr Val Trp Asp Thr Val Gln Ser Ile Asp
122 65 70 75 80
123 Lys Val Lys Gln Glu Met Lys Asn Glu Leu Glu Arg Pro Thr Ile Gln
124 85 90 95
125 Val Phe Leu Thr Asp Leu Phe Gln Asn Asp Phe Asn Ser Val Phe Met
126 100 105 110
127 Leu Leu Pro Ser Phe Tyr Arg Lys Leu Glu Lys Glu Asn Gly Arg Lys
128 115 120 125
129 Ile Gly Ser Cys Leu Ile Ala Ala Met Pro Gly Ser Phe His Gly Arg
130 130 135 140
131 Leu Phe Pro Glu Glu Ser Met His Phe Leu His Ser Ser Tyr Ser Leu
132 145 150 155 160
133 Gln Phe Leu Ser Gln Val Pro Ser Gly Leu Val Thr Glu Leu Gly Ile
134 165 170 175
135 Thr Ala Asn Lys Arg Ser Ile Tyr Ser Ser Lys Ala Ser Pro Pro Pro
136 180 185 190
137 Val Gln Lys Ala Tyr Leu Asp Gln Phe Thr Lys Asp Phe Thr Thr Phe
138 195 200 205
139 Leu Arg Met Arg Ser Glu Glu Leu Leu Ser Arg Gly Arg Met Leu Leu
140 210 215 220
141 Thr Cys Ile Cys Lys Gly Asp Glu Cys Asp Gly Pro Asn Thr Met Asp
142 225 230 235 240
143 Leu Leu Glu Met Ala Ile Asn Asp Leu Val Ala Glu Gly Arg Leu Gly
144 245 250 255
145 Glu Glu Lys Leu Asp Ser Phe Asn Val Pro Ile Tyr Thr Ala Ser Val
146 260 265 270
147 Glu Glu Val Lys Cys Met Val Glu Glu Gly Ser Phe Glu Ile Leu
148 275 280 285
149 Tyr Leu Gln Thr Phe Lys Leu Arg Tyr Asp Ala Gly Phe Ser Ile Asp
150 290 295 300
151 Asp Asp Cys Gln Val Arg Ser His Ser Pro Val Tyr Ser Asp Glu His
152 305 310 315 320
153 Ala Arg Ala Ala His Val Ala Ser Leu Ile Arg Ser Val Tyr Glu Pro
154 325 330 335
155 Ile Leu Ala Ser His Phe Gly Glu Ala Ile Ile Pro Asp Ile Phe His
156 340 345 350
157 Arg Phe Ala Thr Asn Ala Ala Lys Val Ile Arg Leu Gly Lys Gly Phe
158 355 360 365
159 Tyr Asn Asn Leu Ile Ile Ser Leu Ala Lys Lys Pro Glu Lys Ser Asp
160 370 375 380

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/971,020

DATE: 01/02/2002

TIME: 11:08:40

Input Set : A:\026350-068.ST25.txt

Output Set: N:\CRF3\01022002\I971020.raw

```

161 Ile
162 385
165 <210> SEQ ID NO: 4
166 <211> LENGTH: 1360
167 <212> TYPE: DNA
168 <213> ORGANISM: Coffea arabica
170 <400> SEQUENCE: 4
171 gtcctgcata tgaatggagc tccaagaagt cctgcatatg aatggaggcg aaggcgaagc 60
172 aagctacgcc aagaattcat ccttcaatca actggttctc gccaaagtga aacctgtcct 120
173 tgaacaatgc gtacgggaat tggtgcgggc caacttgccc aacatcaaca agtgcattaa 180
174 agttgcagat ttgggatgcg cttccggacc aaacacactt ttaaccgttt gggacactgt 240
175 acaaaagtatt gacaaagtta agcaagaaat gaagaatgaa ttagaacgtc ccaccattca 300
176 ggtttttctg actgatcttt tccaaaatga tttcaattcg gttttcatgc tctgccaag 360
177 cttctaccgc aaacttgaga aagaaaatgg acgcaaaata ggatcgtgcc taatagccgc 420
178 aatgcctggc tctttccacg gcagactctt ccccgaggag tccatgcatt tttacactc 480
179 ttcttacagt cttcagtttt tatcccaggt tcccagcggg ttggtgactg aattggggat 540
180 cactgcgaac aaaaggagca ttactcttc caaagcaagt cctccgcccg tccagaaggc 600
181 atatttgat caatttacga aagattttac cacattttta aggatgcgtt cggaagagtt 660
182 gctttcacgt ggccgaatgc tcttacttg catttgtaaa ggagatgaat gcgacggccc 720
183 gaataccatg gacttacttg agatggcaat aaacgacttg gttgctgagg gacgtctggg 780
184 ggaagaaaaa ttggacagtt tcaatgttcc aatctataca gtttcagtag aagaagtaaa 840
185 gtgcatgggt gaggaggaag gttcttttga aattttatac ttgcagactt ttaagctccg 900
186 ttatgatgct ggcttctcta ttgatgatga ttgccaaagta agatcccatt ccccagtata 960
187 cagcgatgaa catgctagag cagcgcatgt ggcatcatta attagatcag tttacgaacc 1020
188 catcctagca agtcattttg gagaagctat tatacctgac atattccaca ggtttgcgac 1080
189 gaatgcagca aaggttatcc gcttgggcaa aggtctctat aataatctta tcatttctct 1140
190 tgccaaaaaa ccagagaagt cagacatata aaagcttggt tttagttggt tttgtgtta 1200
191 tgggttggtt tctgatacgg ggaaggatt cagtgcgggt ggggttctat ccgagtattg 1260
192 tactttttat attattagtt ggtgtataat tattatgta cattgttata ttcgtaataa 1320
193 aagtgcgta caaaaataaa atattttcat aaaaaaaaaa 1360
195 <210> SEQ ID NO: 5
196 <211> LENGTH: 385
197 <212> TYPE: PRT
198 <213> ORGANISM: Coffea arabica
200 <400> SEQUENCE: 5
201 Met Glu Leu Gln Glu Val Leu His Met Asn Gly Gly Glu Gly Asp Ala
202 1 5 10 15
203 Ser Tyr Ala Lys Asn Ser Ser Phe Asn Gln Leu Val Leu Ala Lys Val
204 20 25 30
205 Lys Pro Val Leu Glu Gln Cys Val Gly Glu Leu Leu Arg Ala Asn Leu
206 35 40 45
207 Pro Asn Ile Asn Lys Cys Ile Lys Val Ala Asp Leu Gly Cys Ala Ser
208 50 55 60
209 Gly Pro Asn Thr Leu Leu Thr Val Arg Asp Ile Val Gln Ser Ile Asp
210 65 70 75 80
211 Asp Val Arg Gln Glu Met Lys Asn Glu Leu Glu Arg Pro Thr Ile Gln
212 85 90 95
213 Val Phe Leu Thr Asp Leu Phe Gln Asn Asp Phe Asn Ser Val Phe Met
214 100 105 110

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/971,020

DATE: 01/02/2002

TIME: 11:08:40

Input Set : A:\026350-068.ST25.txt

Output Set: N:\CRF3\01022002\I971020.raw

```

215 Leu Leu Pro Ser Phe Tyr Arg Lys Leu Glu Lys Glu Asn Gly Arg Lys
216      115      120      125
217 Ile Gly Ser Cys Leu Ile Ala Ala Met Pro Gly Ser Phe His Gly Arg
218      130      135      140
219 Leu Phe Pro Glu Glu Ser Met His Phe Leu His Ser Ser Tyr Ser Leu
220 145      150      155      160
221 Gln Phe Leu Ser Gln Val Pro Ser Gly Leu Val Thr Glu Leu Gly Ile
222      165      170      175
223 Thr Ala Asn Lys Arg Ser Ile Tyr Ser Ser Lys Ala Ser Pro Pro Pro
224      180      185      190
225 Val Gln Lys Ala Tyr Leu Asp Gln Phe Thr Lys Asp Phe Thr Thr Phe
226      195      200      205
227 Leu Arg Ile Arg Ser Glu Glu Leu Leu Ser Arg Gly Arg Met Leu Leu
228      210      215      220
229 Thr Cys Ile Cys Lys Gly Asp Glu Phe Asp Gly Pro Asn Thr Met Asp
230 225      230      235      240
231 Leu Leu Glu Met Ala Ile Asn Asp Leu Val Val Glu Gly His Leu Glu
232      245      250      255
233 Glu Glu Lys Leu Asp Ser Phe Asn Val Pro Ile Tyr Ala Ala Ser Val
234      260      265      270
235 Glu Glu Leu Lys Cys Ile Val Glu Glu Glu Gly Ser Phe Glu Ile Leu
236      275      280      285
237 Tyr Leu Glu Thr Phe Lys Leu Arg Tyr Asp Ala Gly Phe Ser Ile Asp
238      290      295      300
239 Asp Asp Cys Gln Val Arg Ser His Ser Pro Glu Tyr Ser Asp Glu His
240 305      310      315      320
241 Ala Arg Ala Ala His Val Ala Ser Leu Leu Arg Ser Val Tyr Glu Pro
242      325      330      335
243 Ile Leu Ala Asn His Phe Gly Glu Ala Ile Ile Pro Asp Ile Phe His
244      340      345      350
245 Arg Phe Ala Thr Asn Ala Ala Lys Val Ile Arg Leu Gly Lys Gly Phe
246      355      360      365
247 Tyr Asn Asn Leu Ile Ile Ser Leu Ala Lys Lys Pro Glu Lys Ser Asp
248      370      375      380
249 Ile
250 385
253 <210> SEQ ID NO: 6
254 <211> LENGTH: 1304
255 <212> TYPE: DNA
256 <213> ORGANISM: Coffea arabica
258 <400> SEQUENCE: 6
259 ttttagcagtc ccaattcgat ttatgtacaa gtctgcata tgaatggagc tccaagaagt 60
260 cctgcatatg aatggaggcg aaggcgatgc aagctacgcc aagaattcat cctticaatca 120
261 actggttctc gccaaagtga aacctgtcct tgaacaatgc gtaggggaat tgttgcgggc 180
262 caacttgccc aacatcaaca agtgcattaa agttgcggat ttgggatgcg ctccggacc 240
263 aaacacactt ttaacagttc gggacattgt acaaagtatt gacaaagtta ggcaagaaat 300
264 gaagaatgaa ttagaacgtc ccaccattca ggtttttctg actgatcttt tccaaaatga 360
265 tttcaattcg gttttcatgt tgctgccaaag tttctaccgc aaacttgaga aagaaaatgg 420
266 acgcaagata ggatcgtgcc taatagccgc aatgcctggc tctttccacg gcagactctt 480

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/971,020

DATE: 01/02/2002

TIME: 11:08:41

Input Set : A:\026350-068.ST25.txt

Output Set: N:\CRF3\01022002\I971020.raw

L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date